

Life's Solution

Inevitable Humans in a Lonely Universe

SIMON CONWAY MORRIS

University of Cambridge



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

<http://www.cambridge.org>

© Simon Conway Morris 2003

This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2003

Printed in the United Kingdom at the University Press, Cambridge

Typeface Trump Mediaeval 9.5/13.5 pt *System* L^AT_EX 2_ε [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

Conway Morris, S. (Simon)

Life's solution: inevitable humans in a lonely universe / Simon Conway Morris.
p. cm.

ISBN 0 521 82704 3

1. Evolution (Biology) – Philosophy. 2. Convergence (Biology) I. Title.

QH360.5 .C66 2003

576.8'01–dc21 2002035069

ISBN 0 521 82704 3 hardback

Contents

<i>Preface. The Cambridge sandwich</i>	<i>page xi</i>
<i>Acknowledgements</i>	<i>xvii</i>
<i>List of abbreviations</i>	<i>xix</i>
1. Looking for Easter Island	1
Inherency: where is the ground plan in evolution?	5
The navigation of protein hyperspace	8
The game of life	10
Eerie perfection	13
Finding Easter Island	19
2. Can we break the great code?	22
The ground floor	23
DNA: the strangest of all molecules?	27
3. Universal goo: life as a cosmic principle?	32
A Martini the size of the Pacific	33
Goo from the sky	34
Back to deep space	38
A life-saving rain?	42
4. The origin of life: straining the soup or our credulity?	44
Finding its path	47
Problems with experiments	49
On the flat	53
Back to the test tube	58
A sceptic's charter	63
5. Uniquely lucky? The strangeness of Earth	69
The shattered orb	69
Battering the Earth	71
The Mars express	75

	Making the Solar System	77
	Rare Moon	87
	Just the right size	92
	Jupiter and the comets	93
	Just the right place	99
	A cosmic fluke?	105
6.	Converging on the extreme	106
	Universal chlorophyll?	106
	The wheels of life?	111
	Fortean bladders	112
	A silken convergence	115
	Matrices and skeletons	117
	Play it again!	120
	Attacking convergence	126
	Convergence: on the ground, above the ground, under the ground	134
7.	Seeing convergence	147
	A balancing act	148
	Returning the gaze	151
	Eyes of an alien?	158
	Clarity and colour vision	166
	Universal rhodopsin?	170
	Smelling convergence	173
	The echo of convergence	181
	Shocking convergence	182
	Hearing convergence	190
	Thinking convergence?	194
8.	Alien convergences?	197
	Down in the farm	198
	Military convergence	200
	Convergent complexities	205
	Hearts and minds	214
	Honorary mammals	218
	Giving birth to convergence	220
	Warming to convergence, singing of convergence, chewing convergence	223

9.	The non-prevalence of humanoids?	229
	Interstellar nervous systems?	230
	The conceptualizing pancake	231
	The bricks and mortar of life	234
	Genes and networks	237
	Jack, the railway baboon	242
	Giant brains	243
	Grasping convergence	261
	Converging on the humanoid	264
	Converging on the ultimate	271
10.	Evolution bound: the ubiquity of convergence	283
	Ubiquitous convergence	284
	Respiratory convergence	287
	Freezing convergence, photosynthetic convergence	290
	The molecules converge	295
	Convergence and evolution	298
	Converging trends	304
	A possible research programme	308
11.	Towards a theology of evolution	311
	An evolutionary embedment	312
	Darwin's priesthood	314
	Heresy! Heresy!!	318
	Genetic fundamentalism	323
	A path to recovery?	326
	Converging on convergence	328
12.	Last word	331
	<i>Notes</i>	333
	<i>Index</i>	446

I **Looking for Easter Island**

I am a bipedal hominid, of average cranial capacity, write my manuscripts with a fountain pen, and loathe jogging. Thanks to years of work by innumerable biologists I, or anyone else, can tell you to a fair degree of accuracy when the ability to walk upright began, the rate at which our brain increased to its present and seemingly astonishing size, and the origin of the five-fingered forelimb whose present versatility allows me to hold a pen, not to mention the fishy origin of those lungs that make such a noise as the joggers pass me early in the morning on Cambridge's Midsummer Common.

It is obvious that the entire fabric of evolution is imprinted on and through our bodies, from the architecture of our bony skeleton, to the proteins carrying the oxygen surging through our arteries, and our eyes that even unaided can see at least two million years into the past – the amount of time it has taken for the light to travel from the Andromeda Galaxy. In every case – whether for hand or brain – we can trace an ancestry that extends backwards for hundreds of millions, if not billions, of years. Yet, for all that, both the processes and the implications of organic evolution remain controversial. Now at first sight this is rather odd, because it is not immediately clear what is being called into question. Certainly not the fact of evolution, at least as a historical narrative: very crudely, first bacteria, then dinosaurs, now humans. More specifically in terms of process, Darwin's formulation of the mechanisms of evolution is not only straightforward, but seemingly irrefutable. Organisms live in a real world, and evolve to fit their environment by a process of continuous adaptation. This is achieved by a constant winnowing through the operation of natural selection that scrutinizes the available variation to confer reproductive success on those that, by one yardstick or another, are fitter in the struggle for survival.

So is that all there is to say? The recipe for evolution just given is a decidedly bald summary. One intuitively senses that it is an inherently feeble response to an extraordinarily rich history that has brought forth an immense coruscation of form and diversity. Among

living forms this ranges across many scales of complexity, from bacteria that build colonies like miniature trees¹ to immense societies of ants whose populations run into the millions and, independent of us, have stumbled across the advantages of agriculture (Chapter 8). And it is a history that is by no means confined to the complexity of colonies or the limpidity of a geometric shell. It is as much in the range, scope, and acuity of living organisms. They may be mere machines, but consider those owls whose hearing can pinpoint within a two-degree arc the rustling made by a mouse,² the navigational abilities of albatrosses across the seemingly trackless Southern Ocean³ (Fig. 1.1), or even Nellie the cat that smelled Madagascar across more than two hundred miles of ocean.⁴ But despite our admiration, wonder, and – if we are candid – even awe, surely we can still offer the following paraphrase: evolution happens, this bone evolved from that one, this molecule from that one. To be sure, not every transformation and transition will be elucidated, but we are confident this is because of a lack of information rather than a failure of the method.

Yet despite the reality that, as it happens, we humans evolved from apes rather than, say, lizards, let alone tulips, the interpretations surrounding the brute fact of evolution remain contentious, controversial, fractious, and acrimonious. Why should this be so? The heart of the problem, I believe, is to explain how it might be that we, a product of evolution, possess an overwhelming sense of purpose and moral identity yet arose by processes that were seemingly without meaning. If, however, we can begin to demonstrate that organic evolution contains deeper structures and potentialities, if not inevitabilities, then perhaps we can begin to move away from the dreary materialism of much current thinking with its agenda of a world now open to limitless manipulation. Nor need this counter-attack be anti-scientific: far from it. First, evolution may simply be a fact, yet it is in need of continuous interpretation. The study of evolution surely retains its fascination, not because it offers a universal explanation, even though this may appeal to fundamentalists (of all persuasions), but because evolution is both riven with ambiguities and, paradoxically, is also rich in implications. In my opinion the sure sign of the right road is a limitless prospect of deeper knowledge: what was once baffling is now clear, what seemed absurdly important is now simply childish, yet still the journey is unfinished.

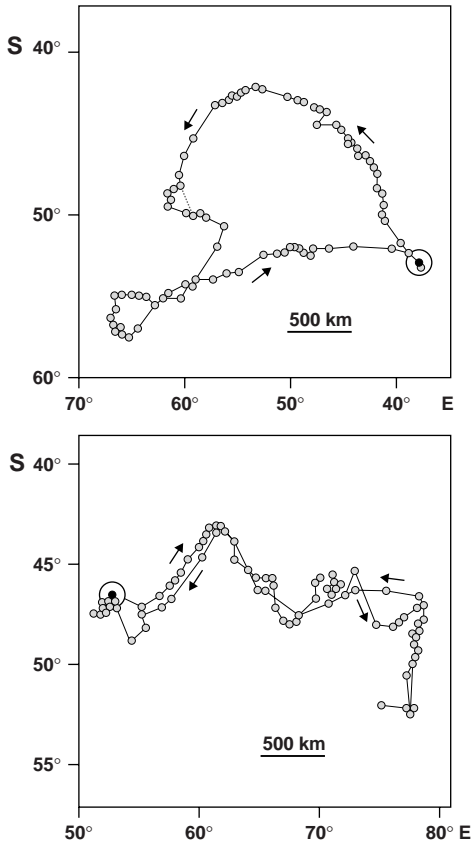


FIGURE 1.1 Two trackways, obtained by satellite monitoring, of the Wandering Albatross across the Southern Ocean. Dots indicate data intercepts, and arrows direction of travel. The upper panel is a departure from South Georgia, on its 13-day trip it passed the Falkland Islands and subsequently Tierra del Fuego. Apart from the distance covered, note the near-straight-line intercept for home. Lower panel is an excursion from Crozet Islands; note how close are the outward and return pathways. (Redrawn from fig. 4b of P. A. Prince *et al.* (1992), Satellite tracking of wandering albatrosses (*Diomedea exulans*) in the South Atlantic, *Antarctic Science*, vol. 4, pp. 31–6 (upper panel) and fig. 8A of H. Weimerskirch *et al.* (1993), Foraging strategy of Wandering Albatrosses through the breeding season: A study using satellite telemetry, *The Auk*, vol. 110, pp. 325–42 (lower panel), with permission of the authors, Cambridge University Press, and *The Auk*.)

One such ambiguity is how life itself may have originated. As we shall see (in Chapter 4) there is no reason to doubt that it occurred by natural means, but despite the necessary simplicity of the process, the details remain strangely elusive. Life itself is underpinned by a rather simple array of building blocks. Most notable are the four (or more accurately five) nucleotides (that is molecules, such as adenine, consisting of a ring of carbon atoms with an attached nitrogen and a sugar) that comprise the DNA (and RNA). The other key building blocks are twenty-odd amino acids that when arranged in chains form the polypeptides and ultimately the proteins. Yet, from this, by various elaborations, has arisen the immense diversity of life. At first sight this would seem to encapsulate the entire process of evolution, yet it soon becomes clear that we hardly understand in any detail the links between the molecular substrate and the nature of the organism. To be sure, there is some crude correlation between the total number of genes and the complexity of the organism, but when we learn that the 'worm' of molecular biologists (the nematode *Caenorhabditis elegans*), which has a relatively simple body plan with a fixed number of cells, has more genes than the 'fly' (the fruit-fly *Drosophila*) with its complex form and behaviour, then there should be pause for thought (see Chapter 9).

One response is to reconsider what we mean by 'the gene'. In particular, it is time to move away from a crippling atomistic portrayal and rethink our views. As has been pointed out by numerous workers, the concept of the gene is without meaning unless it is put into the context of what it is coding *for*, not least an extremely sophisticated biochemistry. Nor are these the only complications. It is well known that significant quantities of DNA, at least in the eukaryotic cell (that is a cell with a defined nucleus and organelles such as mitochondria), are never employed in the process of coding. Pejoratively labelled as 'junk DNA' or 'parasitic DNA', it may be just that, silent and surplus DNA churned out by repeated rounds of duplication of genetic material, like an assembly line commandeered by lunatic robots.⁵ Such a view fits well with the notion that evolution is a process of blind stupidity, a meaningless trek from primordial pond to glassy oceans dying beneath a swollen Sun.

So, beyond the brute fact that evolution happens, the mechanisms and the consequences remain the subject of the liveliest debate and not infrequently acrimony. But, contrary to the desires and beliefs

of creation 'science', the reality of evolution as a historical process is not in dispute. And whatever the divergences of opinion, which as often as not have a tacit ideological agenda concerning the origins of human uniqueness, there is a uniform consensus that vitalism was safely buried many years ago, and the slight shaking of the earth above the grave marking the resting place of teleology is certainly an optical illusion. But is it an illusion? Perhaps as the roots and the branches of the Tree of Life are more fully explored our perspectives will begin to shift. Evolution is manifestly true, but that does not necessarily mean we should take it for granted: the end results, be it the immense complexity of a biochemical system or the fluid grace of a living organism, are genuinely awe-inspiring. Could it be that attempts to reinstall or reinject notions of awe and wonder are not simply delusions of some deracinated super-ape, but rather reopen the portals to our finding a metaphysic for evolution? And this in turn might at last allow a conversation with religious sensibilities rather than the more characteristic response of either howling abuse or lofty condescension.

INHERENCY: WHERE IS THE GROUND PLAN IN EVOLUTION?

Although much of this book will be concerned with retelling the minutiae of biological detail in support of the general thesis of the ubiquity of evolutionary convergence and, what is more important, its implications, here is a brief overview of what strike me as the basic tensions in evolution. The first is what, for want of a better name, I might term 'inherency'. A hard-boiled reductionist will dismiss this as a non-problem, but I am not so sure. Perhaps the first obvious clue was the result, surprising at the time, of the minimal genetic difference between ourselves and the chimps. In terms of structural genes the much-quoted difference amounts, it is said, to about 0.4%. If there were any residual doubt of the closeness between *Homo* and *Pan*, then other indicators of similarity, such as the fact that the string of amino acids that make up the protein haemoglobin is identical in number and sequence, are surely a sufficient indicator of our evolutionary proximity. This, of course, confirms the obvious: we and the chimps share an ancestor, probably between about 6 and 12 million years ago, and indeed there is much we have in common. But in other respects we are poles apart. I'm told that chimps driving cars (or at least go-karts) have the time of their lives, but we are neither likely to see a chimp

designing a car, nor for that matter mixing the driest of Martinis, let alone being haunted by existentialist doubts.

This problem of inherency, however, is far more prevalent and pervasive than the local quirk that chimps and humans are genomically almost identical, but otherwise separated by an immense gulf of differences. Let us look, for example, at a much deeper stage in our evolution, effectively at the time of the ancestors of the fish. Enter the moderately undistinguished animal known as the lancelet worm or amphioxus (*Branchiostoma* and its relatives, Fig. 1.2). By general agreement this beast is the nearest living approximation to the stage in evolution that preceded the fish, which in turn clambered on to land, moved to using the egg, grew fur, and in one lineage developed into socially alert arborealists. All these changes and shifts must have been accompanied by genetic changes, but if we look back to amphioxus we see a genetic architecture in place that seemingly has no obvious counterpart in its anatomy. To give just one example: the central nervous system of amphioxus is really rather simple. It consists of an elongate nerve cord stretching back along the body, above the precursor of the vertebral column (our backbone, consisting of a row of vertebrae) and a so-called brain. The brain can only be described as a disappointment. It is little more than an anterior swelling (it is called the cerebral vesicle) and has no obvious sign in terms of its morphology of even the beginnings of the characteristic threefold division seen in the vertebrate brain of hind-, mid- and fore-sections. Yet the molecular evidence,⁶ which is also backed up by some exquisitely fine studies of microanatomy,⁷ suggests that, cryptically, the brain of amphioxus has regions equivalent to the tripartite division seen in the vertebrates.

The clear implication of this is that folded within the seemingly simple brain of amphioxus is what can almost be described as a template for the equivalent organ of the vertebrates: in some sense amphioxus carries the inherent potential for intelligence. Quite how the more complex brain emerges is yet to be established. The evidence that a key development in the molecular architecture of the vertebrates was episodes of gene duplication,⁸ that is, doubling up of a gene, could well give one clue. This is because the 'surplus' gene is then potentially available for some new function. It could alternatively be claimed that amphioxus is secondarily simplified (the condition sometimes referred to as *regressive*), but it retained genes for

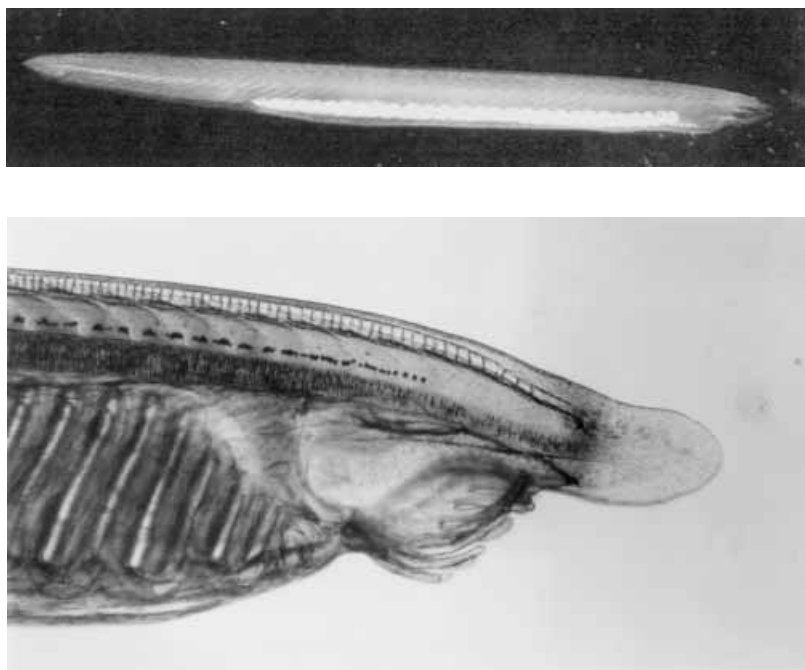


FIGURE 1.2 The amphioxus animal. Upper, entire animal. The anterior end is to the right, with the 'brain' located towards upper side. Prominent white units are gonads. Lower, detail of anterior with prominent feeding (buccal) tentacles and more posteriorly gill bars. The notochord is the longitudinal structure slightly above the mid-line, with closely spaced vertical lines. The nerve cord lies above the notochord, with minimal enlargement at the anterior. (Courtesy of Dale Stokes, Scripps Institution of Oceanography (upper) and Thurston Lacalli, University of Victoria, British Columbia (lower).)

vital functions, although ones no longer specifically connected to the coding for a complex brain. Unfortunately the rather limited information on the earliest amphioxus-like animals, from the Cambrian period (c. 545–500 Ma (million years) ago)⁹ does not extend to seeing their brains. In general, however, the genomic evidence suggests that the living amphioxus is not in some sense degenerate but is genuinely primitive.

Revealing the foundations of the molecular architecture that underpins our brains and sentience gives us not only a feeling of emergence, but underlines how little we really know about why and how organic complexity arises. Nor is this example of the amphioxus brain

and its molecular inherency in any way unusual. Equally instructive examples can be culled from the most primitive animals, such as the sponges and *Hydra* (the latter is a relative of the sea-anemones and corals), in which genes (or proteins) that are essential for complex activities in more advanced animals are already present. Doubtless they have their functions, but what these are and how they have been re-deployed, co-opted, or realigned in more advanced animals is for the most part still unknown. The unravelling of these evolutionary stories is going to be one of the most fascinating episodes in recent biological history, but what will almost certainly be more extraordinary is how much of organic complexity will be seen to be latent in more primitive organisms. Or perhaps not that extraordinary: it is sometimes forgotten that the main principle of evolution, beyond selection and adaptation, is the drawing of new plans but relying on the tried and trusted building blocks of organic architecture.

THE NAVIGATION OF PROTEIN HYPERSPACE

Life, then, is full of inherencies. We might legitimately enquire whether there is anything in the human condition that could prefigure some future level of complexity that with the virtue of hindsight will, no doubt, seem to have been inevitable, but to us remains unimaginable. Yet whatever privileges exist for untrammelled speculation, there is a story to be told which will occupy the rest of this book. My critics will, I imagine, complain at its eclectic, if not unorthodox, nature; and given that the topics covered will include such matters as extraterrestrial amino acids and ants pursuing warfare it is advisable to try to explain the underlying thread of the argument. Here we can do no better than to look at a stimulating and thoughtful essay written by Temple Smith and Harold Morowitz,¹⁰ which is an exploration of the tension between the predictabilities of physical systems and the seemingly contingent muddle that we call history. In brief, and their paper contains many other insights, they remind us of the simplicity of the basic building blocks of life, by which is meant such molecules as amino acids (which go to build the proteins, perhaps familiar as collagen or haemoglobin), or sugars (which when joined together can form carbohydrates). In the case of amino acids, however, even with the rather modest total of the 20 available variants and taking a relatively simple protein – consisting, say, of 100 such building blocks – it is immediately apparent that the potential number of combinations in

which this protein could be assembled is absolutely gigantic. Specifically it is 100^{20} , which is equivalent to 10^{39} .

This is an uncomfortably large number,¹¹ and, as Smith and Morowitz see it, this immensity of possibilities confers an inherent unpredictability on the process of evolution. Taking the figures given above, that is the 20 available amino acids and their random inclusion in a protein composed of a total of 100 amino acids, Smith and Morowitz then apply some apparently stringent criteria to the selection of those proteins that will actually work. The specific function they have in mind for proteins is as natural catalysts, that is, to function as the organic molecules known as enzymes,¹² which serve to accelerate metabolic processes. The alternative, of course, is that a hypothetical protein will be non-functional, failing in one way or another. We know that in principle this is perfectly feasible, because there are many examples known where only a handful of changes, and sometimes even the substitution of a single amino acid for another one, will render the entire protein inoperative and thereby biologically useless. Let us then suppose that only one in a million proteins will be soluble, a necessary prerequisite for the watery milieu of a cell. Let us further suppose, and again the figure seems reasonable, that of these again only one in a million has a configuration suitable for it to be chemically active. How many potentially enzymatically active soluble proteins with an amino acid length totalling 100 could we expect to be available to life? A few thousand, perhaps even a few million? In fact, the total far exceeds the number of stars in the universe.

As Smith and Morowitz dryly note, 'It is quite clear from such numerology that the domain of possible organisms is enormously large if not infinite',¹³ especially when we recall that many proteins are substantially longer than 100 amino acids. The only way we can begin to envisage such a protein domain is in the abstract terms of a kind of hyperspace. Mathematically this will encompass all the measurements that together serve to define the totality of this 'protein space'. As Smith and Morowitz point out, with such an immense number of potential possibilities the number of proteins known to exist on Earth can only be an infinitesimally small fraction of this vast total. As they say, notwithstanding 'the immensity of the dimensionality of the descriptive hyperspace', the world we know and the evolutionary processes that define it have 'produced a very sparsely sampled hyperspace in the actual living world'.¹⁴

One inference that might be immediately drawn from this is that in principle the likelihood of any other world employing an area of 'protein space' that is even remotely close to that found on Earth should be vanishingly small. At this early stage of the argument we can leave aside, for the moment, the distinct likelihood that Earth-like planets are going to be in exceedingly short supply (Chapter 5), and simply remind ourselves that even as our net of exploration spreads first across the Milky Way and then from galaxy to galaxy, so each time a protein chemist steps on to the surface of a new planet only another tiny fraction of this immense 'hyperspace' will be documented. The combinatorial possibilities are so much more immense than all the planets with all their biospheres that most proteins will for ever be only hypothetical constructs. That, at least, is the expectation and it would seem difficult to refute. All other worlds might be expected to be truly alien, at least in so far as the occupation of protein 'space' is concerned. That is, at least, the assumption.

THE GAME OF LIFE

Nevertheless, despite Smith and Morowitz's calculus of immensity, matters are probably much less alien than might at first be imagined. This is because at one level the strings of amino acids and their exact sequence are irrelevant, so long as the protein works effectively. To be sure, specific regions of a protein may be exceedingly sensitive to which amino acid is present, but we also know that various proteins have evolved independently of each other to perform a similar, if not identical, function. Such examples of molecular convergences are examined in more detail later (Chapter 10). What matters here is that these convergences emphatically do not depend on arriving at a closely similar sequence of amino acids, which given the size of protein 'hyperspace' would be almost a miracle. Navigation through this 'hyperspace' depends rather on two principal properties that, as it happens, underpin all life. The first concerns the remarkable specificities of particular sites within the protein that confer the necessary function, for example in those microbial pathogens whose existence depends on precise molecular mimicry to outwit a host's defences. The secondary property is that the complex functions that characterize proteins depend not only, in many cases, on highly specific sites, but also on particular architectural forms that are highly recurrent.¹⁵ As we shall see in at least some protein designs, such as those that render tissue transparent (as in an eye lens), transport or store oxygen (for

respiration), or are sensitive to light (rhodopsin and cryptochromes), the same solution to the biological need has been arrived at independently several times.

The implications are far-reaching, because the 'nodes' of biological possibility may, because of physical constraints, be much more limited than is usually supposed. As Michael Denton and Craig Marshall remark,

If forms as complex as the protein folds are intrinsic features of nature, might some of the higher architecture of life also be determined by physical law? The robustness of certain cytoplasmic forms . . . suggests that [they] may also represent uniquely stable and energetically favoured structures . . . If it does turn out that a substantial amount of higher biological form is natural, then the implications will be radical and far-reaching. It will mean that physical laws must have had a far greater role in the evolution of biological form than is generally assumed. And it will mean . . . that underlying all the diversity of life is a finite set of natural forms that will recur over and over again anywhere in the cosmos where there is carbon-based life.¹⁶

I agree. Not all is possible, options are limited, and different starting points converge repeatedly on the same destinations. Any such evolutionary journey, including navigation through protein 'hyperspace' must presuppose intermediary stages. And here there may be further constraints because seemingly 'sensible' paths may turn out to be non-functional.¹⁷ The 'landscape' of biological form, be it at the level of proteins, organisms, or social systems, may in principle be almost infinitely rich, but in reality the number of 'roads' through it may be much, much more restricted.

This is not to say that there are no alternatives: patently there are, and the world is a diverse place. Smith and Morowitz remind us that despite these potential immensities the actual 'Game of Life', as they call it, is still going to be played the same way everywhere. Here are the four basic rules, which incidentally presuppose variation (which is offered by the different alleles (determining characters) of a gene) and subsequent process of selection.

- (1) Hindsight and foresight are strictly forbidden. Of course we are fully entitled to hug ourselves with delight as we trace, for example, the multiple evolutionary origins of the electric organs

in certain fish (a topic I return to in Chapter 7), but we can only retrodict and not predict.

- (2) Minor changes are easier than major changes. That's something all biologists recognize,¹⁸ and why, for example, there is a deep-seated distrust of macroevolutionary 'jumps' that allow a fully fledged body plan to emerge from some strikingly dissimilar ancestor.
- (3) Resources are not unlimited: the world is finite, and ultimately energy and space are in restricted supply.
- (4) Life has no option but to carry on; it must always play the best hand it can no matter how poor and disastrous the hand might be, and no matter who or what offers the challenge.

In this way Smith and Morowitz neatly encapsulate what evolution is all about. They suggest that given these four basic rules for the Game of Life we should not be surprised to see the emergence both of evolutionary trends and of emerging complexity; Smith and Morowitz also remind us that symbioses¹⁹ and sex are two good ways of speeding up the game. And that is all there is to it? Not quite. Trying to keep the surprise out of their joint voices, Smith and Morowitz then continue:

There is at least one major evolutionary trend not immediately explained by our strategy rules [i.e. their Game of Life]. That being the numerous examples of morphological convergence. Why, in the sparsely sampled genetic space, have there been so many cases of apparent convergence or parallelism? It is surprising in the light of the high probability for novelty to find, even in similar niches, high morphological similarity in distinctly different genetic lines . . . there may be additional rules operating at coarser levels of the genetic space which are less statistical than those discussed.²⁰

This prescient statement prefigures the main purpose of this book, that evolution is indeed constrained, if not bound. Despite the immensity of biological hyperspace I shall argue that nearly all of it must remain for ever empty, not because our chance drunken walk failed to wander into one domain rather than another but because the door could never open, the road was never there, the possibilities were from the beginning for ever unavailable. This implies that we may not only be on the verge of glimpsing a deeper structure to life, but that

it matters little what our starting point may have been: the different routes will not prevent a convergence to similar ends.²¹

EERIE PERFECTION

The understanding of the genetic code was, after the elucidation of the structure of DNA with its four bases and famous double helix, the next triumph in the field of molecular biology. As already noted, proteins are built from the twenty available amino acids,²² although it has long been known that particular examples, such as the protein collagen that goes to form such structures as tendons (Achilles' heel) or the silk proteins that form the spider's web, are enriched in particular amino acids which reflect, in ways that even now are not completely understood, the functional and structural properties of these and other proteins. Thus collagens are enriched in such amino acids as proline, while spider-silks possess notable quantities of alanine and glycine.

Each of the amino acids is coded for by a set of three nucleotide base pairs, accordingly known as a triplet. The original code is, of course, stored in the DNA of the chromosomes, but the actual synthesis of the amino acids occurs through the agency of the RNA in minute structures within the cell known as the ribosomes. Thus, in RNA the four bases are adenine (A), cytosine (C), guanine (G), and uracil (U), the last of which substitutes for thymine (T), which is found in DNA only. With a triplet code and four base pairs there are of course 64 possible combinations. This implies that with only 20 amino acids there is a considerable degree of redundancy, even with the assignment of certain codons to signal 'Start' and 'Stop'. In fact we see that only two amino acids (methionine (abbreviated M) and tryptophan (W)) rely on a single codon each (respectively coded for by AUG and UGG), whereas the remaining 18 amino acids are able to call upon from two to six codons. (For example, histidine (H) uses either CAC or CAU; arginine (R) employs CGU, CGC, CGA, CGG, AGA, and AGG.) It has long been known that this redundancy means that mistakes in coding may not be detrimental; if a substitution within the codon fails to result in the identical amino acid, it stands a good chance of producing another amino acid with similar properties. Amino acids with similar properties, of which their affinity to or repulsion from water (the property of polarity) is particularly important, also tend to have similar pathways of biosynthesis. Here, too, if errors occur then the mistake need not be lethal. For these and other reasons, therefore, it is clear that the

genetic code is excellently adapted to the needs of reliably providing the amino acids that underpin protein construction.

But how good is good? The rule of thumb in evolution is 'good enough to do the job in most circumstances', but not to waste time building a Rolls-Royce of an organism, or, to put it more flippantly, no supersonic albatrosses. Even so, measuring this 'goodness' for purpose is not so easy: organisms themselves are rubbery, slippery, and pliable and non-invasive techniques of investigation are time-consuming and often difficult. One way to address this problem is to look at the design tolerance of an organism, that is, to see the margins of safety built into such a structure as a bone. A powerful analogy, as Jared Diamond reminds us,²³ is to think of a lift in a prestigious building dedicated to the serious accumulation and worship of money. 'Room for one more', says the lift attendant, before the cage shuts, shoots skywards towards the 59th floor, which it never reaches because at the 48th floor the cable snaps . . . Such instances are, in the absence of malice, mercifully rare because the safety factor of such a lift cable, measured as the ratio between its ultimate capacity and maximum load imposed in normal use, is almost 12 times. The equivalent ratio for a cable in a dumb waiter ascending with its cargo of brown Windsor soup and claret is about five; for a bridge, engineers are content to allow a safety factor of only two. In this last case, however, Henry Petroski reminds us that the safety factors for some modern bridges may in reality be perilously small.²⁴

It is perhaps not surprising that by and large the safety factors adopted by organisms²⁵ are closer to those of the dumb waiter and the bridge. Thus the silk dragline of a spider has a modest safety factor of only 1.5, whereas the factor for the leg bone of a kangaroo hopping through the Australian outback is 3. There is an additional and quite important point that many safety factors may in themselves be sub-optimal – spider silk does snap and kangaroos can break their legs – but the margins of safety are necessarily a compromise between strength and many other vital functions in the organism. Even so, over-design does provide an important safety margin, especially when an organism encounters an unpredictable and rare circumstance. In assessing this and other reasons for such safety margins Carl Gans also makes the point that such tolerances may facilitate the occupation of a hitherto untested adaptive zone.²⁶ One of the examples he gives is the New Zealand parrot known as the kea. This is a

fascinating bird with highly adaptable feeding habits. The kea also has a penchant for trashing cars, and its behavioural characteristics include delinquent gangs of young birds.²⁷ In passing I should also mention that notwithstanding the overwhelming evidence for adaptation and functional demands faced by organisms there remain some examples of structures whose significance still baffles biologists. John Currey gives a nice example in the form of the rostral bone in the snout of Blainville's beaked whale (*Mesoplodon densirostris*).²⁸ As the species name suggests, this bone is incredibly dense, but why? One can speculate that it might be employed in fighting, but this rostral bone is very brittle, a consequence of its very low organic content. Alternatively, it might act as ballast,²⁹ but Currey is candid when he writes, 'At the moment, its function, in this rarely found whale, is a mystery'.³⁰

By this stage you will be wondering what possible connection could exist between the safety factors of a kangaroo, let alone the rostral bone of a rare whale, and the efficiency of the genetic code. The point, simply, is that given the realities of the physical world and adaptation, organisms and their components should be designed to do the job adequately, but no more. Humans shudder at the prospect of hurtling to their doom down a lift shaft, and so incorporate a safety margin that seems to be found very seldom in organisms. And at first sight this is what we should see in the genetic code: it certainly isn't random; in fact it is really rather good. But in recent years a group of molecular biologists, notably Steve Freeland and Laurence Hurst, have been trying to arrive at a more precise answer.³¹

Their approach is computer-based, and the basic aim is to randomize the genetic code and then compare the efficiency of a certain fraction of the vast number of alternative codes the computer can generate with the real one, here on Earth. There is, of course, the implicit assumption that a genetic alphabet composed of two base pairs (that is AT/CG),³² as well as the system of triplet codons and the 20 amino acids³³ available for protein construction found in all terrestrial life represents some sort of norm. Alternatives to codon usage and the number and type of amino acids can, of course, be envisaged, but Arthur Weber and Stanley Miller have gone so far as to suggest that 'If life were to arise on another planet, we would expect that ... about 75% of the amino acids would be the same as on the earth.'³⁴ Naturally we need to be cautious in assuming that even if proteins

are universal they necessarily depend on the terrestrial mechanism of codons³⁵ and the same battery of amino acids. Yet there still may be constraints. Codons built as doublets, i.e. only two base pairs (e.g. AA or AU) to code for an amino acid, would probably be rather vulnerable, while quartet or quintet (e.g. AAAA or AUAUA) codons might be getting cumbersome. There are, of course, many more amino acids known than are actually employed in the proteins and, as we shall see (Chapter 3), some of these are best known from meteorites and have no biological equivalents. Even so, given that the simplest amino acids (such as glycine, serine, and alanine) are probably the most readily synthesized anywhere in the Universe, it is possible that they predispose the biosynthetic pathways that lead to the more complex amino acids.³⁶ So, perhaps both the genetic code and protein construction 'out there' are not so very different.

There is, however, a second difficulty in deciding just how effective the terrestrial code might be. This is because randomizing the existing genetic code leads to an astronomical number of alternative possibilities: Freeland and his co-workers suggest a figure of about 10^{18} , which, as they helpfully remind us, is ten times as many seconds as have elapsed since the formation of the Earth. It is another big number (see note 11), and echoes the point I raised in discussing the essay by Smith and Morowitz (see note 10), that with the immensity of a protein, or in this case, genetic 'hyperspace', it would not only be *a priori* exceedingly unlikely that any two biospheres – separated also by a gulf of many light years – would arrive at the same evolutionary solution, but it would be even more fantastically improbable that the solution achieved was not only good (the process of natural selection should see to that) but in fact the very best. Yet, this appears to be the implication in the work by Freeland and his colleagues.

Their work, as is customary, has proceeded in several stages. Well aware of the preceding work already indicating the general efficiency of the genetic code, they examined a million alternative codes (Figure 1.3). To the first approximation the distribution can be compared to the familiar bell-shaped curve that, it is said, describes the distribution of human intelligence (IQ): a few stupid people and equally few geniuses, with most of us somewhere in the middle. So, too, with the distributions of alternative genetic codes: there is a wide range of efficiencies; some alternatives are extremely inefficient ('disastrous') and, perhaps not surprisingly, the majority are quite efficient but not

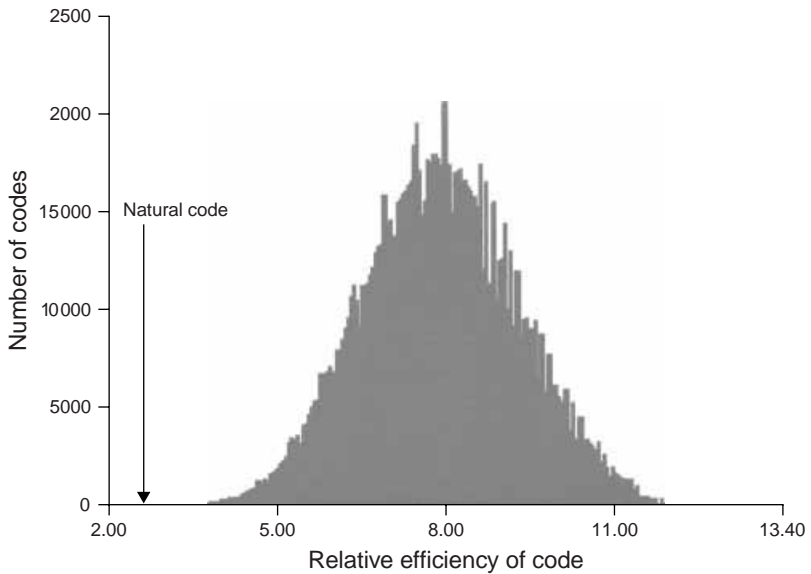


FIGURE 1.3 Eerie perfection. The relative efficiency of randomized genetic codes, ranging from disastrous on the right to increasingly competent to the left. Note the approximately bell-shaped curve: most codes are pretty good, a few terrible, and a few very good. Also note where this planet's genetic code falls: far, far to the left. (Reproduced with permission from *Journal of Molecular Evolution*, from the article 'The genetic code is one in a million', by S.J. Freeland and L.D. Hurst, vol. 47, pp. 238–248, fig.7; 1998, copyright Springer-Verlag, and also with the permission of the authors.)

remarkably effective. Very few of the alternatives are really impressive, but note where in Figure 1.3 the real or natural code falls. Freeland and Hurst have difficulty in keeping the surprise out of their report, even given the proviso that their approach necessitates a number of assumptions. They write: 'the natural genetic code shows *startling* [my emphasis] evidence of optimization, two orders of magnitude higher than has been suggested previously. Though the precise quantification used here may be questioned, the overall result seems fairly clear: under our model, of 1 million random variant codes produced, only 1 was better . . . than the natural code – our genetic code is quite literally "1 in a million"'.³⁷

This result, however, needs to be put into a wider context, because the million (10^6) alternatives that Freeland and Hurst looked at is only a small fraction of the total number of possibilities, which,

as already noted, they estimate to be about 10^{18} . On this basis there could still be an astronomically large number of alternative genetic codes, each of which in its 'local' context could also prove to be very good indeed when compared to a randomly chosen set of a million other codes. In their analysis of the million alternatives Freeland and Hurst specifically noted that the one code that in principle might be better than the natural one had, as one might expect, little similarity to the one used by life on Earth. It seems, however, that the potential figure of 10^{18} alternatives is, in reality, inflated. This is because not all the biosynthetic pathways used to construct the 20 different amino acids are in themselves viable. In a subsequent analysis Freeland and his co-workers suggest that the number of alternative codes that overall are realistically functional is relatively small. They estimate that this number might be about 270 million; and taking into account the similarities between certain amino acids they conclude, again in my opinion startlingly, 'that nature's choice [on Earth] might indeed be the best possible code'.³⁸

In one way we should hardly be surprised at the efficiency of the genetic code.³⁹ It is difficult to believe that the genetic code is not a product of selection, but to arrive at the best of all possible codes selection has to be more than powerful, it has to be overwhelmingly effective. The reason for saying this is that with some minor, and evidently secondary, exceptions,⁴⁰ the genetic code is universal to life: you, the primrose on the table, and the bacteria in your gut all employ the same code. The earliest evidence for life is about 3.8 billion years ago and these forms are presumably directly ancestral to all groups still alive today. If so, this indicates that whatever changes occurred as the genetic code evolved towards its stable state must have been achieved still earlier; the genetic code would not otherwise be universal. Yet, as we shall see (Chapter 4), life itself may not be older than about 4 billion years. Two hundred million years (and possibly much less) to navigate to the best of all possible codes, or at least from the 270 million alternatives? Part of the explanation, as is so often the case in evolution, may be to look for a step-like arrangement: once one stage is achieved, other things then become so much more likely.⁴¹ Yet, there is also a sense that given a world of DNA and amino acids, then perhaps the genetic code we know is more or less an inevitable outcome. And if this is true, then what else might be inevitable, both here on Earth and elsewhere?

This is not the only way to look at inevitabilities in evolution. The argument from the genetic code looks to a potentially gigantic 'hyperspace' of alternative possibilities, yet the evidence suggests that rapidly and with extraordinary effectiveness a very good, perhaps even the best, code is arrived at. It is as if the Blind Watchmaker takes off her sunglasses and decides to visit her brother Chronos. Off she sets, crossing streets roaring with traffic driven by psychotics, through the entrails of the subway system of a megalopolis, and, after catching a series of intercontinental express trains with connection times of two minutes each, she arrives at Chronos' front door at 4 p.m. prompt, just in time for a relaxing cup of tea.

FINDING EASTER ISLAND

Hence to an explanation for the title of this first chapter. Easter Island is the remotest speck of land on Earth, surrounded by the vastness of the Pacific Ocean. At first sight it seems quite extraordinary that it could have been encountered by the seafaring Polynesians, however audacious. Surely, one would suppose, it was a chance discovery, perhaps by mariners who had been blown far off course, which led to the prows of the first canoes accidentally grating onto a beach of Easter Island perhaps some 1500 years ago. Another quirk of history? Very probably not. Easter Island may have marked one of the furthest points in this great human diaspora, but its discovery was inevitable given the sophisticated search strategy of the Polynesians. As Geoffrey Irwin has shown,⁴² not only were these people superb navigators, but they developed a method of quartering the ocean that aimed to find new lands. Century by century their net of exploration widened. When a particular season failed in the objective they had a sure way of finding their way home to safety. Their vessels were designed for protracted journeys, but the key to their success was to head against the prevailing winds on the outward journey. At the limit of their range on any one journey, the sternward winds rapidly returned them towards their home and safety. And how was home, another speck in the ocean, arrived at? In the sky above the boats the net of stars provided the clues to celestial navigation, and as the constellations fitted into place so an increasingly familiar starlit sky provided the beacons for a successful homecoming.

So, too, in evolution. Isolated 'islands' provide havens of biological possibility in an ocean of maladaptedness (Fig. 1.4). No wonder

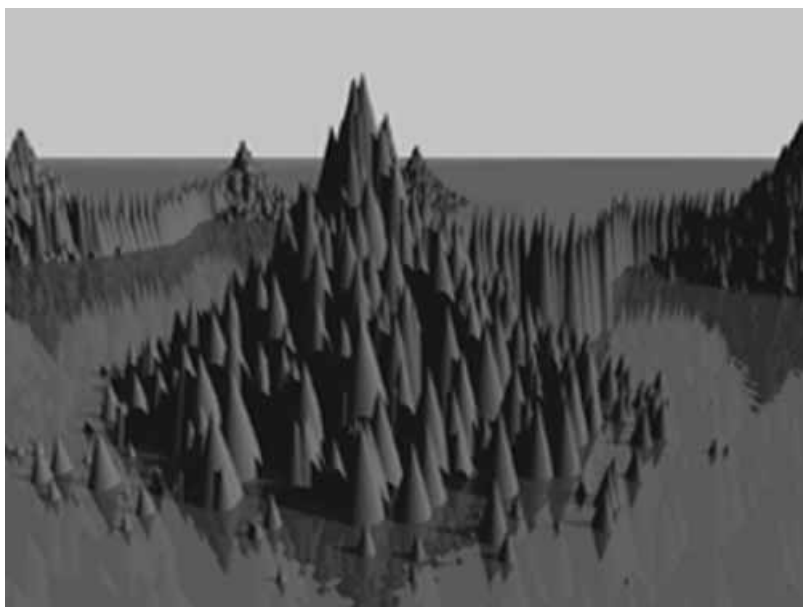


FIGURE 1.4 A metaphorical view of protein 'hyperspace', in which functional proteins project above an immense 'ocean' that submerges non-functional alternatives. (Reprinted from *Journal of Molecular Biology*, vol. 301, D.D. Axe, Extreme functional sensitivity to conservative amino acid changes on enzyme exteriors, pp. 585–596, fig. 5. Copyright 2000, with permission from Elsevier Science, and also with the permission of the author.)

the arguments for design and intelligent planning have such a perennial appeal. Whether it be by navigation across the hyperdimensional vastness of protein space, the journey to a genetic code of almost eerie efficiency, or the more familiar examples of superb adaptation, life has an extraordinary propensity for its metaphorical hand to fit the glove. Life depends both on a suitable chemistry, whose origins are literally cosmic, and on the realities of evolutionary adaptation. The chemistry is acknowledged but largely ignored; the adaptation is often derided as a wishful fantasy. As with the audacious and intelligent Polynesians, so life shows a kind of homing instinct. Its central paradox revolves around the fact that despite its fecundity and baroque richness life is also strongly constrained. The net result is a genuine creation, almost unimaginably rich and beautiful, but one also with an underlying structure in which, given enough time, the inevitable must happen.